

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-092766

(43)Date of publication of application : 06.04.2001

(51)Int.Cl.

G06F 13/00

(21)Application number : 11-266018

(71)Applicant : HITACHI LTD

(22)Date of filing : 20.09.1999

(72)Inventor : YAMAGUCHI TATSUYA
AIZAWA YOSHIICHI

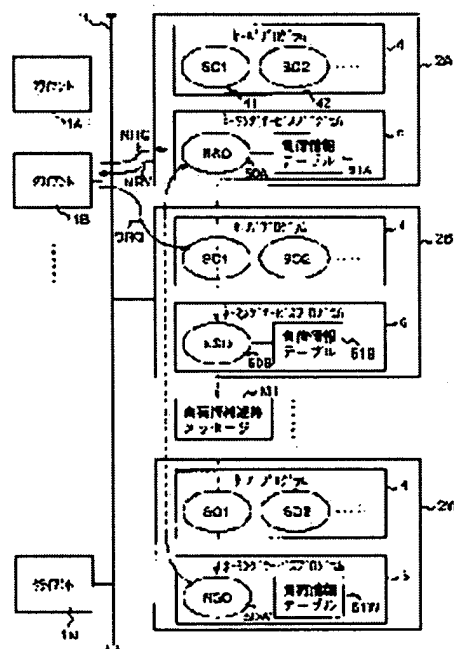
(54) CLIENT SERVER SYSTEM AND NAMING SERVICE METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a dynamic load distribution naming service suitable to a client server system having many clients.

SOLUTION: In the client server system consisting of plural server machines 2 to be connected with the clients 1 via a network 3, respective server machines 2 are provided with plural kinds of server objects 40 to offer information service to the clients and naming service objects 50 to specify the server objects to execute the information service to the clients. The respective naming service objects are provided with routines 505 to collect load information of other server machines and executes dynamic naming service according to load states of the respective server machines to client terminals.

図 1



(1)

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-92766

(P2001-92766A)

(43) 公開日 平成13年4月6日 (2001.4.6)

(51) Int.Cl.

G 0 6 F 13/00

識別記号

3 5 7

F I

G 0 6 F 13/00

テーマコード(参考)

3 5 7 Z 5 B 0 8 9

審査請求 未請求 請求項の数10 O L (全 20 頁)

(21) 出願番号

特願平11-266018

(22) 出願日

平成11年9月20日 (1999.9.20)

(71) 出願人

000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者

山口 達也

神奈川県横浜市戸塚区戸塚町5030番地 株

式会社日立製作所ソフトウェア事業部内

(72) 発明者

相澤 宣一

神奈川県横浜市戸塚区戸塚町5030番地 株

式会社日立製作所ソフトウェア事業部内

(74) 代理人

100068504

弁理士 小川 勝男 (外1名)

Fターム(参考) 5B089 GA11 GA21 GB09 JA11 KA06

KB06 KB11 KC27 KH03 MA03

MC06

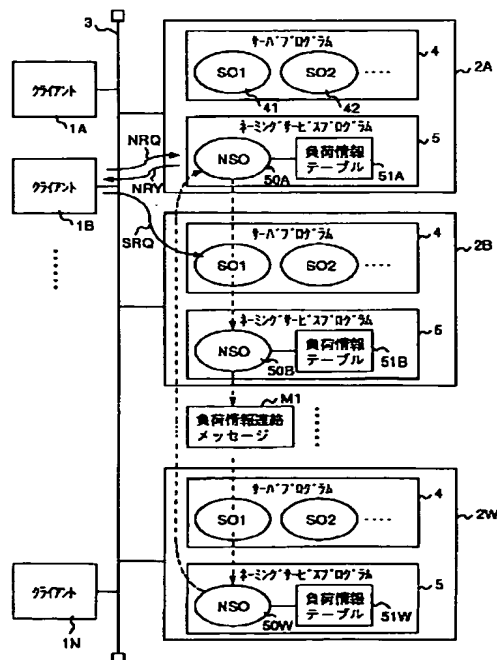
(54) 【発明の名称】 クライアントサーバシステムおよびネーミングサービス方法

(57) 【要約】

【課題】 多数のクライアントをもつクライアントサーバシステムに適した動的負荷分散ネーミングサービスを提供する。

【解決手段】 ネットワーク3を介してクライアント1と接続される複数のサーバマシン2からなるクライアントサーバシステムにおいて、各サーバマシン2が、クライアントに対して情報サービスを提供するための複数種類のサーバオブジェクト40と、クライアントに情報サービスを実行すべきサーバオブジェクトを指定するためのネーミングサービスオブジェクト50とを備える。各ネーミングサービスオブジェクトは、他のサーバマシンの負荷情報を収集するためのルーチン505を有し、クライアント端末に対して各サーバマシンの負荷状態に応じた動的なネーミングサービスを実行する。

図 1



【特許請求の範囲】

【請求項 1】 ネットワークを介してクライアント端末と接続される複数のサーバマシンからなるクライアントサーバシステムにおいて、

上記各サーバマシンが、クライアント端末に対して情報サービスを提供するための複数種類のサーバオブジェクトと、クライアント端末に情報サービスを実行すべきサーバオブジェクトを指定するためのネーミングサービスオブジェクトとを備え、

上記各ネーミングサービスオブジェクトが、他のサーバマシンの負荷情報を収集するための手段を有し、クライアント端末に対して各サーバマシンの負荷状態に応じた動的なネーミングサービスを実行することを特徴とするクライアントサーバシステム。

【請求項 2】 前記各ネーミングサービスオブジェクトが、他のサーバマシンと負荷情報を交信するため手段と、収集された複数のサーバマシンの負荷情報を記憶するための負荷情報テーブルと、上記負荷情報テーブルを参照して低負荷状態にあるサーバマシンから順に前記クライアント端末からの要求に合致したサーバオブジェクトの有無を判定し、上記クライアント端末に割り当てるべきサーバオブジェクトを決定するための手段とを有することを特徴とする請求項 1 に記載のクライアントサーバシステム。

【請求項 3】 前記負荷情報テーブルが、前記各サーバマシンの負荷情報と障害の有無を記憶し、前記サーバオブジェクト決定手段が、上記負荷情報テーブルを参照して、障害状態にあるサーバマシンを除外して、前記クライアント端末に割り当てるべきサーバオブジェクトを決定することを特徴とする請求項 2 に記載のクライアントサーバシステム。

【請求項 4】 前記複数のサーバマシンのうちの 1 つが、負荷情報連絡メッセージを生成するマスタ用ネーミングサービスオブジェクトを有し、前記各交信手段が、上記マスタ用ネーミングサービスオブジェクトで生成した負荷情報連絡メッセージを前記複数のネーミングサービスオブジェクトに循環させるように順次に次のネーミングサービスオブジェクトに転送し、上記負荷情報連絡メッセージを介して各サーバマシンの負荷情報を交信することを特徴とする請求項 2 または請求項 3 に記載のクライアントサーバシステム。

【請求項 5】 前記複数のサーバマシンが複数のグループに分割され、各グループ毎に前記マスタ用ネーミングサービスオブジェクトを有し、前記各交信手段が、上記マスタ用ネーミングサービスオブジェクトで生成された負荷情報連絡メッセージをグループ内で循環させるように順次に次のネーミングサービスオブジェクトに転送し、上記負荷情報連絡メッセージを介して各サーバマシンの負荷情報を交信することを特徴とする請求項 2 または請求項 3 に記載のクライアントサーバシステム。

【請求項 6】 前記複数のマスタ用ネーミングサービスオブジェクトのうちの 1 つが、グループ間連絡メッセージを生成するための手段を有し、前記マスタ用ネーミングサービスオブジェクトにおける各交信手段が、上記グループ間連絡メッセージを順次に他のマスタ用ネーミングサービスオブジェクトに転送し、上記メッセージを介してグループ間でサーバマシンの負荷情報を交信することを特徴とする請求項 5 に記載のクライアントサーバシステム。

【請求項 7】 ネットワークを介してクライアント端末と接続される複数のサーバマシンからなり、上記各サーバマシンが、クライアント端末に対して情報サービスを提供するための複数種類のサーバオブジェクトと、クライアント端末に情報サービスを実行すべきサーバオブジェクトを指定するためのネーミングサービスオブジェクトとを備えたクライアントサーバシステムにおけるネーミングサービス方法において、

上記ネーミングサービスオブジェクトのうち、マスタに指定されたネーミングサービスオブジェクトが、上記複数のサーバマシンを一巡させる負荷情報連絡メッセージを生成し、該負荷情報連絡メッセージに自サーバマシンの新たな負荷情報を設定して上記ネットワークに送出するステップと、

上記負荷情報連絡メッセージの宛先となったネーミングサービスオブジェクトが、上記負荷情報連絡メッセージから他のサーバマシンの負荷情報を取得すると共に、該負荷情報連絡メッセージに自サーバマシンの新たな負荷情報を設定して、次のサーバマシンのネーミングサービスオブジェクト宛に送出するステップと、

クライアント端末からネーミングサービス要求を受けたネーミングサービスオブジェクトが、低負荷状態にあるサーバマシンから順に上記要求に合致したサーバオブジェクトの有無を判定し、上記クライアント端末にサーバオブジェクトを割り当てるステップとからなることを特徴とするネーミングサービス方法。

【請求項 8】 前記マスタに指定されたネーミングサービスオブジェクトが、前記負荷情報連絡メッセージが所定期間で循環するように、前記複数のサーバマシンを一巡した負荷情報連絡メッセージを次の送信タイミングとなるまで保持することを特徴とする請求項 7 に記載のネーミングサービス方法。

【請求項 9】 前記マスタに指定されたネーミングサービスオブジェクトが、前記負荷情報連絡メッセージが連続的に循環するように、前記複数のサーバマシンを一巡した負荷情報連絡メッセージから他のサーバマシンの負荷情報を取得し、該負荷情報連絡メッセージに自サーバマシンの新たな負荷情報を設定した後、次のサーバマシンのネーミングサービスオブジェクト宛に送出することを特徴とする請求項 7 に記載のネーミングサービス方法。

【請求項 10】 前記各ネーミングサービスオブジェクト

が、送出した負荷情報連絡メッセージに対して所定時間内に受信応答がなかった場合、上記負荷情報連絡メッセージに前記宛先ネーミングサービスオブジェクトが障害となったことを示す情報を設定し、一巡ルート上で上記障害ネーミングサービスオブジェクトの次に位置するネーミングサービスオブジェクト宛に送出し、クライアント端末からネーミングサービス要求を受けたネーミングサービスオブジェクトが、上記負荷情報連絡メッセージで通知された障害ネーミングサービスオブジェクトを備えたサーバマシンを除外して、前記サーバオブジェクトの割り当て動作を行うことを特徴とする請求項7～請求項9の何れかに記載のネーミングサービス方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、クライアントサーバシステムおよびネーミングサービス方法に関し、更に詳しくは、同一の情報サービスを2以上のサーバマシンから選択的に提供できるように、それぞれが複数種類のサーバプログラム（サーバオブジェクト）を備えた複数のサーバマシンからなるクライアントサーバシステムにおいて、サーバ負荷を分散するように各クライアント端末へのサーバオブジェクトの割り当てを動的に行えるようにしたクライアントサーバシステムおよびネーミングサービス方法に関する。

【0002】

【従来の技術】同一機能のサーバプログラム（サーバオブジェクト）を複数のサーバマシンに重複して装備しておき、クライアント端末（以下、単にクライアントと言う）に対して上記何れかのサーバマシンから情報サービスを提供するように構成したクライアントサーバシステムにおいては、多数のクライアントからのサービス要求が特定のサーバあるいは少数のサーバ群に集中したことによってクライアントへのサービス性能が低下しないように、できるだけサーバ負荷を分散させた形で、クライアントに対するサーバあるいはサーバオブジェクトの割り当て（以下、ネーミングと言う）を行う必要がある。

【0003】負荷分散を考慮したネーミングサービス方法として、例えば、特開平7-93238号公報（従来技術1）には、クライアントからネームサーバにサーバアドレスの獲得を要求し、ネームサーバが、複数のサーバの中から負荷の低いサーバを選択し、そのアドレスをクライアントに通知することによって動的に負荷分散するサーバ選択方式が提案されている。また、特開平9-319689号公報（従来技術2）には、クライアント側にダイナミックリンクライブラリ（DLL）を設けておき、クライアントプログラムからサーバ名を指定されたとき、上記DLLが、ネームサーバから、負荷の低い順にサーバの選択順序を指定したアドレスリストを入手し、該アドレスリストに従って各サーバにプローブパケ

ットを送出し、その応答を判定することによって、上記サーバ名で特定される所望のサービスを提供中の低負荷サーバを選択するようにしたサーバ選択方式が提案されている。

【0004】

【発明が解決しようとする課題】然るに、上記従来技術1、2では、クライアントに情報サービスを提供する複数のサーバマシン以外に、これらのサーバの負荷状態を監視しながらネーミングサービスを行うネーミング専用のサーバを必要としている。このように専用のネームサーバを設置したクライアントサーバシステムでは、ネームサーバに問い合わせが集中するため、クライアント数の増加に伴ってネーミングサービスが低下すると言う問題がある。また、従来技術2では、不特定多数のクライアント端末にそれぞれ特殊機能をもつDLLを装備する必要がある。

【0005】本発明の目的は、ネーミングサービスを高速化できるクライアントサーバシステムおよびネーミングサービス方法を提供することにある。本発明の他の目的は、サーバマシンの負荷分散を動的に行い、ネーミングサービスを高速化できるクライアントサーバシステムおよびネーミングサービス方法を提供することにある。本発明の更に他の目的は、サーバマシン台数とクライアント数が増加しても、サーバ負荷情報を効率的に収集してネーミングサービスに反映できるクライアントサーバシステムおよびネーミングサービス方法を提供することにある。

【0006】

【課題を解決するための手段】上記目的を達成するために、本発明のクライアントサーバシステムは、ネットワークを介してクライアント端末と接続される複数のサーバマシンからなり、各サーバマシンが、クライアント端末に対して情報サービスを提供するための複数種類のサーバオブジェクトと、クライアント端末に情報サービスを実行すべきサーバオブジェクトを指定するためのネーミングサービスオブジェクトとを備え、上記各ネーミングサービスオブジェクトが、他のサーバマシンの負荷情報を収集するための手段を有し、クライアント端末に対して各サーバマシンの負荷状態に応じた動的なネーミングサービスを実行することの特徴とする。上記構成によれば、ネーミングサービス機能が情報サービス用の複数のサーバマシンに分散して配置されているため、サーバシステムに接続されるクライアント数が増加した場合でも、ネーミングサービス要求を複数のサーバマシンに分散させることによって、迅速な応答が可能となる。

【0007】更に詳述すると、本発明のクライアントサーバは、上記各ネーミングサービスオブジェクトが、他のサーバマシンと負荷情報を交信するため手段と、収集された複数のサーバマシンの負荷情報を記憶するための負荷情報テーブルと、上記負荷情報テーブルを参照して

低負荷状態にあるサーバマシンから順に前記クライアント端末からの要求に合致したサーバオブジェクトの有無を判定し、上記クライアント端末に割り当てるべきサーバオブジェクトを決定するための手段とを有することを特徴とする。本発明の1実施例によれば、上記複数のサーバマシンのうちの1つが、負荷情報連絡メッセージを生成するマスタ用ネーミングサービスオブジェクトを有し、上記各交信手段が、上記マスタ用ネーミングサービスオブジェクトで生成した負荷情報連絡メッセージを複数のネーミングサービスオブジェクトに循環させるように順次に次のネーミングサービスオブジェクトに転送し、上記負荷情報連絡メッセージを介して各サーバマシンの負荷情報を交信する。

【0008】上記構成によれば、各ネーミングサービスオブジェクトが、上記負荷情報連絡メッセージから他のサーバマシンの負荷情報を取得し、該負荷情報連絡メッセージに自サーバマシンの新たな負荷情報を設定して次のネーミングサービスオブジェクトに転送することによって、負荷情報連絡メッセージの受信の都度、他の複数のサーバマシンの負荷状態を把握することが可能となる。また、送出した負荷情報連絡メッセージに対して所定時間内に受信応答がなかった場合、上記負荷情報連絡メッセージに宛先ネーミングサービスオブジェクトが障害となったことを示す情報を設定し、一巡ルート上で上記障害ネーミングサービスオブジェクトの次に位置するネーミングサービスオブジェクト宛に送出するようにすれば、サーバマシンの負荷情報と共に障害情報も各ネーミングサービスオブジェクトに通知できる。この場合、上記障害情報を負荷情報テーブルに記憶しておくことによって、クライアント端末からネーミングサービス要求を受けた時、各ネーミングサービスオブジェクトが、障害ネーミングサービスオブジェクトと対応するサーバマシンを除外して、サーバオブジェクトの割り当て動作を行うようにすることができる。

【0009】尚、マスタに指定されたネーミングサービスオブジェクトに、上記複数のサーバマシンを一巡した負荷情報連絡メッセージを次の送信タイミングとなるまで保持させれば、負荷情報連絡メッセージを所定周期で循環させることができる。また、マスタに指定されたネーミングサービスオブジェクトが、上記複数のサーバマシンを一巡した負荷情報連絡メッセージから他のサーバマシンの負荷情報を取得し、該負荷情報連絡メッセージに自サーバマシンの新たな負荷情報を設定した後、直ちに次のサーバマシンのネーミングサービスオブジェクト宛に送出するよう動作させれば、負荷情報連絡メッセージを連続的に循環させることができる。

【0010】前者の場合は、負荷情報連絡メッセージの巡回回数が少なくなるため、ネットワーク上の通信負荷と負荷情報テーブルを更新するためのオーバーヘッドを低下させることが可能となる。後者の場合、負荷情報連

絡メッセージの巡回頻度が高くなるため、負荷情報が頻繁に更新され、各ネーミングサービスオブジェクトは、他のサーバマシンの最新の負荷情報に基いたネーミングサービスが可能となる。

【0011】サーバマシン台数の多い大規模なクライアントサーバシステムの場合、本発明では、サーバマシンを複数のグループに分割し、各グループ毎にマスタ用ネーミングサービスオブジェクトを指定する。この場合、各交信手段は、上記マスタ用ネーミングサービスオブジェクトで生成された負荷情報連絡メッセージをグループ内で循環させるように順次に次のネーミングサービスオブジェクトに転送し、上記負荷情報連絡メッセージを介して各サーバマシンの負荷情報を交信する。

【0012】上記大規模クライアントサーバシステムにおいて、各ネーミングサービスオブジェクトに、他のグループに属したサーバマシンの負荷情報も考慮してネーミングサービスを実行させたい場合、本発明では、複数のマスタ用ネーミングサービスオブジェクトのうちの1つに、グループ間連絡メッセージの生成機能を与え、マスタ用ネーミングサービスオブジェクトにおける各交信手段が、上記グループ間連絡メッセージを順次に他のマスタ用ネーミングサービスオブジェクトに転送し、上記メッセージを介してグループ間でサーバマシンの負荷情報を交信する。

【0013】上記構成によれば、各マスタ用ネーミングサービスオブジェクトは、上記グループ間連絡メッセージによって、他グループの各サーバマシンの負荷情報を収集し、自グループの各サーバマシンの負荷情報を他グループに通知できる。従って、グループ間連絡メッセージによって収集した他グループの各サーバマシンの負荷情報を前述した負荷情報連絡メッセージに設定し、自グループ内で循環させることによって、自グループの各ネーミングサービスオブジェクトにシステム全体のサーバマシンの負荷情報を配布することが可能となる。

【0014】本発明のネーミングサービス方法は、ネットワークを介してクライアント端末と接続される複数のサーバマシンからなり、上記各サーバマシンが、クライアント端末に対して情報サービスを提供するための複数種類のサーバオブジェクトと、クライアント端末に情報サービスを実行すべきサーバオブジェクトを指定するためのネーミングサービスオブジェクトとを備えたクライアントサーバシステムにおいて、(A) 上記ネーミングサービスオブジェクトのうち、マスタに指定されたネーミングサービスオブジェクトが、上記複数のサーバマシンを一巡させる負荷情報連絡メッセージを生成し、該負荷情報連絡メッセージに自サーバマシンの新たな負荷情報を設定して上記ネットワークに送出するステップと、

(B) 上記負荷情報連絡メッセージの宛先となったネーミングサービスオブジェクトが、上記負荷情報連絡メッセージから他のサーバマシンの負荷情報を取得すると共

に、該負荷情報連絡メッセージに自サーバマシンの新たな負荷情報を設定して、次のサーバマシンのネーミングサービスオブジェクト宛に送出するステップと、(C)クライアント端末からネーミングサービス要求を受けたネーミングサービスオブジェクトが、低負荷状態にあるサーバマシンから順に上記要求に合致したサーバオブジェクトの有無を判定し、上記クライアント端末にサーバオブジェクトを割り当てるステップとからなることを特徴とする。

【0015】

【発明の実施の形態】以下、本発明の実施例を、図面を用いて詳細に説明する。本発明のクライアントサーバシステムは、図1に示すように、高速通信リンク3で相互接続された複数のクライアント端末1(1A、1B、…1N)と、複数のサーバマシン2(2A2、B2、…W)からなる。サーバ間およびサーバとクライアントとの間の通信は、例えば、TCP/IPのような標準的ネットワーク・プロトコルを使用して行われるが、後述するクラスタ間の通信には、TCP/IP以外の特別仕様の軽量プロトコルを適用してもよい。

【0016】本発明のクライアントサーバシステムでは、各サーバマシン2に、サーバプログラム4とネーミングサービスプログラム5とを装備する。上記サーバプログラム4は、クライアントからの要求に応じた各種情報サービスを提供するための複数のプログラム(以下、サーバオブジェクト：SOと言う)40(40-1、40-2、…)からなる。また、上記ネーミングサービスプログラム5は、後述するように複数のサーバ間で相互に負荷情報を交信しながら負荷情報テーブル51を定期的に更新する負荷情報連絡機能と、クライアントから所望のサーバオブジェクトに関してネーミングサービス要求NRQを受けた時、上記負荷情報テーブル51から軽負荷状態にある最適サーバマシンを検索し、該サーバマシン上で動作する所望サーバオブジェクト40-iのアドレス情報(位置情報とオブジェクトリファレンス)をネーミングメッセージNRYによって回答するネーミングサービス機能とを備えている。以下の説明では、上記負荷情報連絡機能を備えたネーミングサービスプログラム5をネーミングサービスオブジェクト(NSO)50と呼ぶことにする。

【0017】本発明のクライアントサーバシステムでは、上述したように、サーバオブジェクト40が動作する各サーバマシン2にネーミングサービス機能を装備してあるため、各クライアント1は、任意のサーバマシンに対してネーミングサービスを要求でき、従来システムのように特定ネームサーバへのネーミングサービス要求の集中を回避できる。この場合、各クライアント毎に、ネーミングサービス要求の宛先となる特定のネーミングサービスオブジェクト(NSO)を予め指定しておいてもよいし、各クライアントが、ネーミングサービス要求

の都度、複数のネーミングサービスオブジェクトの中からランダムに宛先NSOを選択するようにしてもよい。また、本発明のクライアントサーバシステムでは、各ネーミングサービスオブジェクトが、自サーバを含めて、クライアントサーバシステムを構成する複数のサーバマシンの負荷状態を周期的に収集し、クライアントに対して、低負荷状態にあるサーバマシン上で動作中のサーバオブジェクトのアドレス情報をネーミングメッセージNRYで通知するようになっているため、結果的に、各クライアントからのサービス処理要求SRQが低負荷サーバマシンに対して発行され、サーバマシンの動的負荷分散が実現される。図1に示したクライアントサーバシステムでは、破線で示すように、各サーバマシン2が、受信メッセージを順次に隣接サーバに中継することによって、複数のサーバマシン間で負荷情報連絡メッセージM1が循環的に転送されるようにしている。

【0018】図2は、上記負荷情報連絡メッセージM1のフォーマットの1例を示す。図において、61はIPヘッダ、62はTCPヘッダ、63はデータフィールドであり、負荷情報連絡メッセージM1は、上記データフィールド63に設定される。負荷情報連絡メッセージM1は、メッセージの種類を示す識別子631と、メッセージシーケンス番号632と、メッセージの送信元となったNSOの識別子633と、後続する負荷情報データブロックの個数を示すNSO個数634と、複数の負荷情報データブロック635-1～635-mと、メッセージ(IPパケット)の終了を示すフッタ636とからなる。上記負荷情報データブロック635-1～635-mは、負荷情報連絡メッセージM1の循環経路上に含まれるサーバマシンと対応しており、それぞれNSO識別子635aと、負荷情報フィールド635bとからなっている。

【0019】上記負荷情報連絡メッセージM1は、図3に示すように、クライアントサーバシステムを構成している複数のサーバマシンのうち、予めリスタートマスタとして指定された特定のサーバ上のNOS(リスタートマスタNSO)50Aによって生成される。各サーバマシン上のNOSは、負荷情報連絡メッセージM1を受信すると、負荷情報データブロック635-1～635-mが示す他サーバの負荷情報を負荷情報テーブル51に記憶し、上記受信メッセージ中の自NOS識別子と対応するデータブロックの負荷情報を現在の自サーバマシンの負荷情報に更新した後、これを次データブロックのNOS識別子が示すサーバマシンに転送する。この場合、リスタートマスタNSO50Aが、一巡した負荷情報連絡メッセージM1を次のNSO：50Bに直ちに中継することによって、負荷情報連絡メッセージM1が連続的に循環するようにしてもよいし、リスタートマスタNSO50Aが、一巡した負荷情報連絡メッセージM1を次の送信タイミングまで保持することによって、負荷情報

連絡メッセージM1が定期的に循環するようにしてもよい。

【0020】リスタートマスタNOS50Aが動作するサーバマシンに障害が発生すると、負荷情報連絡メッセージM1の循環が途絶え、正常な他のNSO50B～50Fで負荷情報を収集できなくなるため、リスタートマスタNOS50Aに代わって負荷情報連絡メッセージM1を生成する予備のリスタートマスタNSOを予め指定しておく。予備リスタートマスタNSOには、例えば、リスタートマスタNOS50Aから負荷情報連絡メッセージM1を最初に受信するNSO50Bが選ばれる。予備リスタートマスタNSO50Bは、負荷情報連絡メッセージM1の受信が所定時間以上にわたって途絶えた場合、リスタートマスタNOS50Aに障害が発生したものと判断して、リスタートマスタとしての動作を開始する。

【0021】各サーバマシンまたはNSOは、負荷情報連絡メッセージM1を受信すると、IPヘッダの送信元アドレス(DA)612に対して、受信応答メッセージを返送する。また、各NSOは、上述した負荷情報の収集と更新処理を実行した後、IPヘッダの送信元アドレス(SA)611と宛先アドレス(DA)を書き換えて、負荷情報連絡メッセージM1を次のNSOに転送する。負荷情報連絡メッセージM1を転送した後、所定時間内に受信応答メッセージが受信されなかった場合は、宛先のNSOに障害が発生したものと判断する。例えば、図3において、NSO50DがNSO50Eに負荷情報連絡メッセージM1を転送した後、所定時間内に受信応答メッセージが受信されなかった場合、NSO50Dは、負荷情報連絡メッセージM1内のNSO50Eと対応する負荷情報フィールドに障害コードを設定し、これを新たな負荷情報連絡メッセージM1'として、障害NSO50Eの次に位置するNSO50F宛に転送する。障害NSO50Eと次のNSO50Fとの関係は、負荷情報連絡メッセージM1における負荷情報データブロック635の配列によって決まる。

【0022】上述した負荷情報連絡メッセージM1'の転送によって、障害NSO50E以外の各NSOに、NSO50Eが動作するサーバマシン2Eが障害状態にあることを通知でき、各NSOにサーバマシン2Eを除外したネーミングサービスを実行させることができる。尚、サーバマシン2Eが障害から回復した場合、NSO50Eに、次のNSO50Fに対して、NSO50Eの識別子と負荷情報とを含む再エントリ要求メッセージM3を送信させ、再エントリ要求メッセージM3を受信したNSO50Fに、次の負荷情報連絡メッセージの受信時に、上記NSO50Eの負荷情報データブロックに含まれる障害コードを上記負荷情報に書き換えて、次のNSO50Aに転送するよう動作させる。これによって、NSO50Eの障害回復を他の全てのNSOに通知で

き、NSO50Dから復旧NSO50Eに負荷情報連絡メッセージを中継させることが可能となる。何れかのサーバマシンで負荷情報連絡メッセージM1の受信直後に障害が発生した場合、上記障害サーバマシンにおいて負荷情報連絡メッセージの循環が中断される。この場合、リスタートマスタ59Aが、時間監視によって新たな負荷情報連絡メッセージを生成し、この負荷情報連絡メッセージが循環する過程で、前述したNSO50Eの場合と同様のプロセスで、障害発生個所を検出できる。

【0023】クライアントサーバシステムを構成するサーバマシン2の台数が比較的少数の場合は、比較的短時間で負荷情報連絡メッセージM1を全NSOに循環させることができる。しかしながら、多数のサーバマシンを含む大規模なクライアントサーバシステムでは、負荷情報連絡メッセージM1の循環に時間を要し、各NSOに通知された負荷情報が実際のマシン状態から大幅にずれた値となる可能性がある。本発明では、サーバマシンの台数が増加した場合、図4に示すように、これらのサーバマシンを複数のグループ(クラスタ)G1、G2、G3、…に分割し、各グループ毎にリスタートマスタを指定し、各グループ内で負荷情報連絡メッセージM1を循環させることによって、負荷情報連絡メッセージM1の巡回範囲を局所化し、中継遅延による負荷情報鮮度の劣化を防ぐ。

【0024】上述したようにサーバマシンをクラスタ化した場合、各ネーミングサービスオブジェクトに、各グループ内に限定した負荷分散でネーミングサービスを行わせてもよいが、グループを超えた負荷分散でネーミングサービスを行うためには、各ネーミングサービスオブジェクトに、全グループの負荷情報を含む負荷情報連絡メッセージを巡回させる必要がある。全グループの負荷情報を交信するためには、例えば、リスタートマスタとして指定された複数のNSO(50A、150B、150C)の間で、グループ間連絡メッセージM2を循環させ、上記グループ間連絡メッセージM2を介して、各クラスタ内のサーバ負荷状態を他のクラスタに通知する。この場合、グループ間連絡メッセージM2を生成するために、上記リスタートマスタNSOのうちの1つをグループ間連絡マスタとして予め指定しておく。上記構成によれば、各リスタートマスタが、グループ間連絡メッセージM2で収集した他グループの各サーバマシンの負荷情報を自グループ内で循環する負荷情報連絡メッセージM1に反映させることによって、システム全体のサーバ負荷情報を各ネーミングサービスオブジェクトに通知できる。

【0025】図5は、各サーバマシン2が備えるプログラムの構成を示す。サーバマシン2では、ネットワーク3を介してクライアント1と通信する機能を備えたオペレーションシステム(OS)60と、ネーミングサービスプログラム(ネーミングサービスオブジェクト:NS

O) 5と、複数種類のサーバオブジェクト40-1~40-nを有するサーバプログラム4とが動作している。

【0026】ネーミングサービスオブジェクト(NSO)5は、イニシャライズ処理ルーチン501、振り分け処理ルーチン502、ネーミングサービス処理ルーチン503、時間起動処理ルーチン504、負荷情報連絡処理ルーチン505、時間監視処理ルーチン506、グループ間連絡処理ルーチン507、再エントリ要求処理ルーチン508からなり、これらの複数のコンポーネントルーチンは、マルチプロセス、マルチスレッドで動作する。また、上記ネーミングサービスオブジェクトは、上述したコンポーネントルーチンが参照するテーブルとして、負荷情報テーブル51、イニシャライズ定義ファイル52、再エントリ要求バッファ53、サーバオブジェクト(SO)位置情報テーブル54を備えている。

【0027】図6は、負荷情報テーブル51の構成を示す。負荷情報テーブル51は、クライアント1からネーミングサービス要求を受けた時、ネーミングサービス処理ルーチン503によって参照されるテーブルであり、その内容は、負荷情報連絡処理ルーチン505、グループ間連絡処理ルーチン507、再エントリ要求処理ルーチン508によって更新される。負荷情報テーブル51は、図4で示したクラスタ構成におけるグループG1~G3の識別子を示すグループID510毎に、各グループに所属するネーミングサービスオブジェクトと対応した複数の負荷情報レコードを有し、各負荷情報レコードは、ネーミングサービスオブジェクトを示すNSO識別子511と、上記ネーミングサービスオブジェクトが位置するサーバマシンの負荷状態(CPU利用率)を示す負荷情報512と、サービス割り当て回数65と、フラグ514とを含む。尚、クラスタ化を必要としない比較的小規模なクライアントサーバシステムでは、グループID510は不要である。

【0028】上記グループID510は、グループ連絡マスタを始点としてグループ間連絡メッセージM2の転送順に配列され、NSO識別子511は、各グループ(クラスタ)内のリスタートマスタを始点として負荷情報連絡メッセージM1の転送順に配列されている。上記サービス割り当て回数65は、負荷情報テーブルの更新周期内での負荷の増加を示すための指標として用意されたものであり、上記割り当て回数65の値は、負荷情報テーブルの更新の都度クリアされ、そのサーバマシン上の何れかのサーバオブジェクトをネーミングする毎にインクリメントされる。また、フラグ514は、ネーミングサービス処理ルーチン503が最適サーバマシンを検索する過程でチェック済みのレコードか否かを示すために利用される。

【0029】図7は、サーバマシンが起動された時、イニシャライズ処理ルーチン501によって参照されるイニシャライズ定義ファイル52の内容を示す。イニシャ

ライズ定義ファイル52は、そのネーミングサービスオブジェクトが、負荷情報連絡メッセージM1を生成すべきリスタートマスタか否かを示すリスタートマスタ表示フラグ521と、自グループにおけるリスタートマスタを示すリスタートマスタNSO識別子522と、そのネーミングサービスオブジェクトがリスタートマスタ障害時に負荷情報連絡メッセージM1を生成すべき予備リスタートマスタか否かを示す予備リスタートマスタ表示フラグ523と、負荷情報連絡メッセージM1の転送先となる次のネーミングサービスオブジェクトを示す次NSO識別子524と、そのネーミングサービスオブジェクトがグループ間連絡メッセージを生成すべきグループ連絡マスタか否かを示すグループ連絡マスタ表示フラグ525と、そのネーミングサービスオブジェクトが所属するグループを示すグループ識別子526と、そのネーミングサービスオブジェクトがリスタートマスタまたは予備リスタートマスタに指定されていた場合に、グループ間連絡メッセージM2の転送先となる次のグループのリスタートマスタを示すグループ間連絡メッセージ宛先NSO識別子527と、グループ間連絡メッセージM2の循環ルートにおいて上流側に隣接するリスタートマスタを示すグループ間連絡メッセージ送信元NSO識別子528とを含む。

【0030】図8は、再エントリ要求バッファ53の内容を示す。停止状態にあったサーバマシンが起動された時、または障害状態から回復した時、上記サーバマシン上のネーミングサービスオブジェクトから、稼働中の次のネーミングサービスオブジェクトに対して、再エントリ要求メッセージM3が発行される。再エントリ要求メッセージM3を受信したネーミングサービスオブジェクトは、次に受信した負荷情報連絡メッセージによって、上記再エントリ要求元のサーバマシンが稼働状態になったこと、および、要求元サーバマシンの負荷状態を他のネーミングサービスオブジェクトに通知する。再エントリ要求バッファ53は、次の負荷情報連絡メッセージが生成または受信される迄の間、上記再エントリ要求メッセージの内容を保存しておくためのものであり、再エントリ要求メッセージで通知されたNSO識別子531と負荷情報532とを含む複数のレコードエリアからなっている。再エントリ要求バッファ53の各レコードは、その内容を負荷情報連絡メッセージに反映した時点でクリアされる。

【0031】図9は、サーバオブジェクト(SO)位置情報テーブル54の構成を示す。SO位置情報テーブル54は、各サーバマシン上で動作するサーバオブジェクト40-1~40-nの定義情報を格納するためのものであり、NSO識別子541毎に、サーバオブジェクト名称542と、サーバオブジェクトの位置情報543と、オブジェクトリファレンス544とからなる複数のレコードを記憶している。

【0032】各ネーミングサービスオブジェクト50は、クライアントからネーミングサービス要求を受けると、ネーミングサービス処理ルーチン503が、前述の負荷情報テーブル51から負荷の低い順に選択したNSO識別子511に基づいて上記SO位置情報テーブル54を参照し、ネーミングサービス要求で指定するサービス（サーバオブジェクト名称）が上記NSO識別子511をもつサーバマシンで実行可能か否かを判定する。上記SO位置情報テーブル54上で上記NSO識別子511と対応するレコード領域に指定のサーバオブジェクト名称が登録されていなければ、ネーミングサービス処理ルーチン503は、負荷情報テーブル51から次のNSO識別子を選択し、SO位置情報テーブル54で同様の検索を行い、指定のサーバオブジェクト名称542をもつレコードが見つければ、そのレコードで定義された位置情報543とオブジェクトレファレンス544を要求元のクライアントに回答する。

【0033】図10は、クライアント1からネーミングサービスオブジェクト50に送信されるネーミングサービス要求メッセージNRQのフォーマットを示す。ネーミングサービス要求メッセージNRQは、メッセージ識別子631と、送信元クライアントの位置情報637と、ネーミング対象となるサーバオブジェクト名称638と、フッタ636とからなり、図2に示した負荷情報連絡メッセージと同様、その前部にIPヘッダ61とTCPヘッダ62を付加したIPパケットの形で通信リンク3に送出される。

【0034】図11は、グループ間連絡メッセージM2のフォーマットを示す。グループ間連絡メッセージM2では、付加情報データブロックがグループ毎に配列しており、図2に示した付加情報連絡メッセージM1に含まれる項目631～636以外に、グループ数639とグループ識別子640（640-1、640-2、…）を含む。上記グループ数639の後に、グループ識別子640とNSO個数634に先導されたグループ毎の付加情報データブロック635-1～635-mが配列されている。

【0035】図12は、再エントリ要求メッセージM3のフォーマットを示す。再エントリ要求メッセージM3は、メッセージ識別子631と、該メッセージの発行元ネーミングサービスオブジェクトを示す送信元NSO識別子634と、上記ネーミングサービスオブジェクトが動作しているサーバマシンの付加情報（CPU利用率）635bと、フッタ636とを含む。

【0036】図13は、動作可否確認メッセージM4と、グループ間連絡メッセージ宛先変更メッセージM5のフォーマットを示す。動作可否確認メッセージM4は、他のネーミングサービスオブジェクトが正常に稼動しているか否かを確認するためのメッセージであり、例えば、予備リスタートマスタがリスタートマスタの状態

を確認する場合に発行される。また、グループ間連絡メッセージ宛先変更メッセージM5は、グループ間連絡マスタのサーバマシンに障害が発生した時、これに代わって動作すべき予備リスタートマスタが、インシャライズ定義ファイルのグループ間連絡メッセージ送信元NSO識別子528によって指定されたネーミングサービスオブジェクトに送信するメッセージである。これらのメッセージは、メッセージ識別子631と、送信元NSO識別子634と、フッタ636とからなる。

【0037】図14は、インシャライズ処理ルーチン501のフローチャートを示す。インシャライズ処理ルーチン501では、まず、インシャライズ定義テーブル52のリスタートマスタ表示フラグ521を参照して、自NSOがリスタートマスタに指定されているか否かを判定する（ステップ100）。リスタートマスタに指定されていないければ、再エントリ要求メッセージM3を生成し、次NSO識別子524が示すネーミングサービスオブジェクトに送信し（ステップ101）、振り分け処理ルーチン502を実行する。

【0038】自NSOがリスタートマスタに指定されていた場合、負荷情報テーブル51の各レコードにおいて、負荷情報512、サービス割り当て回数513、フラグ514をクリアした後、サーバマシンの負荷情報（CPU利用率）を取得し、これを負荷情報テーブル51の自NSO識別子と対応するレコードに記録する（102）。次に、自NSO識別子と上記負荷情報とを含み、新たなメッセージシーケンス番号（MSN）632をもつ負荷情報連絡メッセージM1を生成して、次NSO識別子524が示す次のネーミングサービスオブジェクトに送信する（103）。次NSOからの応答メッセージを待ち、一定時間経過しても応答がない場合は、負荷情報連絡メッセージM1の送信リトライ回数が規定回数に達したか否かを判定し（106）、規定回数に達していなければ、ステップ104に戻って、負荷情報連絡メッセージM1の送信動作を繰り返す。次NSOから応答があった場合、または、送信リトライ回数が規定回数に達した場合は、インシャライズ定義テーブル52のグループ間連絡マスタ表示フラグ525をチェックし（110）、もし、自NSOがグループ間連絡マスタに指定されていないければ、このインシャライズ処理ルーチンを終了し、振り分け処理ルーチン502を実行する。

【0039】自NSOがグループ間連絡マスタに指定されていた場合は、自NSO識別子と上記負荷情報とを含み、新たなメッセージシーケンス番号（MSN）632をもつグループ間連絡メッセージM2を生成して、グループ間連絡メッセージ宛先NSO識別子527が示すネーミングサービスオブジェクトに送信する（111）。負荷情報連絡メッセージM1の場合と同様、宛先NSOからの応答メッセージを待ち（112）、一定時間経過しても応答がない場合は、リトライ回数が規定回数に達

する迄、グループ間連絡メッセージM2の送信動作を繰り返す(113、114)。宛先NSOから応答があった場合、または、送信リトライ回数が規定回数に達した場合は、このイニシャライズ処理ルーチンを終了し、振り分け処理ルーチン502を実行する。

【0040】図15は、振り分け処理ルーチン502のフローチャートを示す。振り分け処理ルーチン502では、図16に示す時間監視処理ルーチン506を起動して(ステップ120)、監視タイマをスタートさせておく。この後、入力イベントを待ち(121)、ネーミングサービス要求メッセージNRQを受信した場合は(122)ネーミングサービス処理ルーチン503、時間監視タイマ割り込みが発生した場合は(123)時間起動処理ルーチン504、負荷情報連絡メッセージM1を受信した場合(124)は負荷情報連絡処理ルーチン505、グループ間連絡メッセージM2を受信した場合(125)はグループ間連絡処理ルーチン507、再エントリ要求メッセージM3を受信した場合は(126)再エントリ要求処理ルーチン508をそれぞれ実行し、次の入力イベントを待つ。

【0041】グループ間連絡宛先変更メッセージM5を受信した場合(127)は、受信メッセージの内容に従ってイニシャライズ定義ファイル52を更新し(131)、動作可否確認メッセージM4を受信した場合は、受信応答メッセージを返送する(132)。もし、ネーミングサービスの終了を指示するコマンドが入力された場合は(129)、このルーチンを終了する。

【0042】図16は、時間監視処理ルーチン506のフローチャートを示す。時間監視処理ルーチン506が起動されると、監視タイマをスタートし(ステップ140)、時間監視のリセット要求の発生を待つ(141)。リセット要求が発生した場合は、監視タイマをリセットして(142)、このルーチンを終了する。時間監視のリセット要求がないまま、監視タイマがタイムオーバーとなった場合は(143)、時間起動処理ルーチン504を実行するためのタイマ割り込みを発生して(144)、このルーチンを終了する。

【0043】図17は、負荷情報連絡処理ルーチン505のフローチャートを示す。負荷情報連絡処理ルーチン505では、受信した負荷情報連絡メッセージM1のメッセージシーケンス番号(MSN)631を前回受信したメッセージのMSNと比較し(ステップ150)、受信メッセージが新規メッセージでない場合は、受信メッセージを無視して、このルーチンを終了する。

【0044】受信メッセージが新規メッセージの場合は、時間監視処理ルーチン506にリセット要求を与えた後(151)、データブロックをカウントするためのパラメータiを初期値「1」にする(152)。次に、負荷情報連絡メッセージ内の第i番目のデータブロック635-iに含まれるNSO識別子が負荷情報テーブル

51に登録済みか否かを判定し(153)、未登録の場合は、負荷情報テーブル51に上記データブロック635-iのNSO識別子と負荷情報とを含む新たなレコードを追加し(154)、登録済みの場合は、負荷情報テーブル内の該当レコードの負荷情報512の値を上記データブロック635-iが示す新たな負荷情報に置き換える(155)。この後、パラメータiの値をインクリメントし(156)、パラメータiをNSO個数631の値と比較する(157)。もし、パラメータiがNSO個数を超過していなければ、ステップ153に戻り、負荷情報連絡メッセージ内の次のデータブロックについて、上述した処理を繰り返す。

【0045】パラメータiがNSO個数を越えた場合、再エントリバッファ53に再エントリ要求データが蓄積されているか否かをチェックする(158)。再エントリ要求データが蓄積されていた場合は、再エントリバッファ53から読み出したNSO識別子531と負荷情報532とからなるデータブロックを負荷情報連絡メッセージM1に設定する(159)。この時、負荷情報連絡メッセージM1に上記NSO識別子531と対応するデータブロック635-jが既にあれば、該データブロックの負荷情報635-jの内容を更新し、上記NSO識別子531と対応するデータブロックがない場合は、負荷情報連絡メッセージM1内の自NSO識別子をもつデータブロックの前に、上記再エントリ要求されたNSO識別子をもつデータブロックを追加する。負荷情報連絡メッセージM1への設定を完了した再エントリ要求データは、再エントリ要求バッファから消去する(160)。

【0046】次に、自サーバマシンの負荷情報を取得し、負荷情報テーブル51の自NSO識別子と対応するレコードと上記負荷情報連絡メッセージの該当データブロックに設定(161)した後、イニシャライズ定義テーブル52のリスタートマスタ表示フラグ521をチェックする(162)。自NSOがリスタートマスタに指定されていた場合は、メッセージシーケンス番号(MSN)を更新(163)した後、もし、自NSOがリスタートマスタに指定されていなければ、受信時のMSNのままで、上記負荷情報連絡メッセージM1を次のNSOに転送する。この時、上記メッセージM1のデータブロック配列における自NSO識別子の次に位置したNSO識別子を宛先NSOに選択し、該NSO宛のIPヘッダを生成して、負荷情報連絡メッセージM1を送信し(164、165、166)、宛先NSOからの応答を待つ(167)。一定時間内に応答がなければ、宛先NSOに障害が発生したものと判断する。この場合、負荷情報連絡メッセージM1と負荷情報テーブル51における上記宛先NSOと対応する負荷情報フィールドに障害コードを設定(168)した後、ステップ164に戻る。

【0047】負荷情報フィールドに既に障害コードが設

定されているNSO識別子は宛先NSOの対象外として、メッセージM1のデータブロック配列から更に次のNSO識別子を宛先NSOに選択して、上述した負荷情報連絡メッセージM1の送信動作を繰り返す。これによって、サーバマシンの障害発生に応じて、負荷情報連絡メッセージM1の循環経路を動的に変更しながら、稼動サーバの負荷情報の収集と通知を繰り返すことができる。宛先NSOから一定時間内に応答があった場合、または、メッセージM1のデータブロック配列において選択すべきNSO識別子がなくなった場合は、時間監視処理ルーチン506を起動して(169)、このルーチンを終了する。

【0048】上述したフローチャートでは、リスタートマスタに指定されたネーミングサービスオブジェクトが、他の一般のネーミングサービスオブジェクトと同様、受信メッセージを直ちに次のネーミングサービスオブジェクトに転送するようになっているが、もし、負荷情報連絡メッセージM1を一定に時間間隔において定期的に循環させたい場合は、リスタートマスタに指定されたネーミングサービスオブジェクトが、次の送信タイミングを待つために、ステップ163の代わりに時間監視処理ルーチン506を起動して(169)、このルーチンを終了するようにすればよい。

【0049】図18は、ネーミングサービス処理ルーチン503のフローチャートを示す。ネーミングサービス処理では、負荷情報テーブル51を参照し、フラグ514が“0”状態であり、負荷情報512が最小値となっているレコード(NSO識別子511)を選択する(ステップ170)。負荷情報512に障害コードが設定されたレコードは選択対象外とし、負荷情報テーブル51に上記選択条件に該当するNSO識別子511がない場合は(171)、クライアントにエラーメッセージを送信して(172)、このルーチンを終了する。

【0050】選択条件に合ったレコードが見つかった場合、フラグ514を“1”に設定し(173)、サービス割り当て回数513を予め決められた基準値と比較する(174)。サービス割り当て回数が基準値を超えていた場合は、上記レコードのNSO識別子はビジー状態と判断し、ステップ170に戻って、次のレコード(NSO識別子)を選択する。サービス割り当て回数が基準値以下であれば、SO位置情報テーブル54を参照し、負荷情報テーブル51で選択したNSO識別子511と対応して、ネーミングサービス要求されたサービスオブジェクト名称(SO名)をもつレコードが登録されているか否かをチェックする(175)。SO位置情報テーブル54に該当するレコードがなければ、ステップ170に戻って、次のレコード(NSO識別子)を選択する。

【0051】SO位置情報テーブル54に要求SO名のレコードが登録されていた場合、該レコードで定義され

た位置情報543とオブジェクトレファレンス544に応答メッセージNRYによってクライアントに回答し(177)、負荷情報テーブル51の上記選択されたレコードのサービス割り当て回数513の値をインクリメントする(178)。この後、次の検索処理に備えて、負荷情報テーブル51の全てのフラグ514をクリアし(179)、このルーチンを終了する。

【0052】図19は、時間起動処理ルーチン504のフローチャートを示す。時間起動処理ルーチン504は、前回の負荷情報連絡メッセージM1の送信後、所定時間内に負荷情報連絡メッセージを受信できなかった場合に実行される。また、負荷情報連絡メッセージを定期的に循環させる方式を採用した場合、リスタートマスタにとって負荷情報連絡メッセージM1の次の送信タイミングとなった時、この時間起動処理ルーチン504が実行される。

【0053】時間監視処理ルーチンでは、まず、サーバマシンの負荷情報を取得し、負荷情報テーブル51の自NSO識別子をもつレコードにおいて、負荷情報512の値を更新する(ステップ180)。次に、イニシャライズ定義ファイル52のリスタートマスタ表示フラグを参照し、自NSOがリスタートマスタに指定されているか否かを判定する(181)。自NSOがリスタートマスタに指定されていた場合は、ステップ187で、前回受信済みの負荷情報連絡メッセージのメッセージシーケンス番号(MSN)632の値と、自NSO識別子と対応するデータブロックの負荷情報635とをそれぞれ更新することによって、新たな負荷情報連絡メッセージを生成する(187)。

【0054】自NSOがリスタートマスタに指定されていない場合は、自NSOが予備リスタートマスタに指定されているか否かを判定し(182)、予備リスタートマスタに指定されていないならば、次のNSO識別子のネーミングサービスオブジェクトを宛先とする再エントリ要求メッセージM3を送信(183)した後、時間監視処理ルーチン506を起動して(192)、このルーチンを終了する。自NSOが予備リスタートマスタに指定されていた場合、リスタートマスタNSO識別子522が示すネーミングサービスオブジェクトを宛先とする動作可否確認メッセージM4によって、リスタートマスタの状態を確認する(184)。もし、リスタートマスタから応答があった場合、すなわち、リスタートマスタが正常に動作していた場合は、他のネーミングサービスオブジェクトと同様、ステップ183以降の動作を実行する。

【0055】リスタートマスタから応答がなかった場合、予備リスタートマスタは、イニシャライズ定義ファイル52のリスタートマスタ表示フラグ521をオン状態(“1”)に設定し、グループ間連絡メッセージ宛再NSO識別子528が示すネーミングサービスオブジェ

クトを宛先とするグループ間連絡宛先変更メッセージM5を送信する(186)。この後、予備リスタートマスタは、前述したリスタートマスタと同様の処理動作を行う。すなわち、ステップ187で新たな負荷情報連絡メッセージを生成し、再エントリ要求バッファ53の内容に応じて上記負荷情報連絡メッセージを修正し(188)、該負荷情報連絡メッセージを次のNSO識別子が示すネーミングサービスオブジェクトに送信する(189)。

【0056】リスタートマスタは、更にグループ間連絡マスタ表示フラグ525をチェックし(190)、グループ間連絡マスタに指定されていなければ、時間監視処理ルーチンを起動して(192)、このルーチンを終了する。もし、グループ間連絡マスタに指定されていた場合は、グループ間連絡メッセージ宛先NSO識別子527が示す他グループのネーミングサービスオブジェクトに対して、新たなグループ間連絡メッセージM2の送信し(191)、時間監視処理ルーチンを起動する(192)。上記グループ間連絡メッセージM2は、前回受信したグループ間連絡メッセージM2のメッセージシーケンス番号(MSN)632を更新し、自グループのデータブロック635(635-1、635-2、...)の負荷情報635bを負荷情報テーブル51が示す新たな負荷情報に置き換えたものである。

【0057】図20は、再エントリ要求処理ルーチンのフローチャート508を示す。再エントリ要求処理では、再エントリ要求バッファ53に、再エントリ要求メッセージM3で通知された送信元NSO識別子634と負荷情報635bを追加し(ステップ200)、これらの情報を負荷情報テーブル51にも反映する(201)。上記送信元NSO識別子634に該当するレコードが負荷情報テーブルに既に存在している場合は、負荷情報を書き換え、該当するレコードが負荷情報テーブルにない場合は、自NSO識別子をもつレコードの一つ前に新たなレコードを追加する。

【0058】次に、上記再エントリ要求メッセージM3の送信元NSO識別子634をイニシャライズ定義ファイル52に定義されたリスタートマスタNSO識別子522と比較し(202)、要求メッセージM3の送信元がリスタートマスタでなければ、このルーチンを終了する。要求メッセージM3の送信元がリスタートマスタの場合は、イニシャライズ定義ファイル52のリスタートマスタ表示フラグ521をチェックし(203)、上記フラグがオン状態であれば、これをオフ状態に戻してから(204)、このルーチンを終了する。グループ間連絡メッセージの送信元NSOに対する宛先変更通知は、予備リスタートマスタ側が上記ステップ204で行ってもよいが、再エントリ要求メッセージM3に対する受信応答を受けた時点で、リスタートマスタ側で行ってもよい。

【0059】図21は、グループ間連絡処理ルーチン507のフローチャートを示す。グループ間連絡処理では、受信したグループ間連絡メッセージM2の送信元NSO識別子632とイニシャライズ定義ファイル52に定義されたグループ間連絡メッセージ送信元NSO識別子528とを比較し(ステップ210)、不一致の場合は、イニシャライズ定義ファイルのNSO識別子528を上記受信メッセージの送信元NSO識別子632に置き換える(211)。次に、グループ数をカウントするためのパラメータjの値を初期値(=1)に設定し(212)、受信メッセージM2の第j番目のグループ識別子640-jとイニシャライズ定義ファイル52に定義された自グループ識別子526とを比較する(213)。もし、グループ識別子640-jが自グループ識別子の場合、受信メッセージの第jグループの各負荷情報を負荷情報テーブル51にある自グループの最新の負荷情報で更新(214)した後、パラメータjの値をインクリメントする(218)。

【0060】グループ識別子640-jが自グループ識別子でない場合は、グループ識別子640-jが負荷情報テーブル51に登録済みか否かを判定する(215)。上記グループ識別子640-jが登録済みの場合は、負荷情報テーブル51の上記第jグループと対応するレコードの負荷情報512を上記受信メッセージが示す負荷情報に置き換え(216)、上記グループ識別子640-jが未登録の場合は、負荷情報テーブル51に上記受信メッセージM2の第jグループの各データブロックの内容を追加(217)した後、パラメータjの値をインクリメントする(218)。インクリメントされたパラメータjの値をグループ数639と比較し(219)、パラメータjの値をグループ数を超えていなければ、ステップ213に戻って、次のグループについて同様の動作を繰り返す。パラメータjの値をグループ数を超えた時、すなわち、受信メッセージの全てのグループについて負荷情報テーブルへの反映を終えた時、イニシャライズ定義ファイルのグループ間連絡マスタ表示フラグ525を参照して、自NSOがグループ間連絡マスタに指定されているか否かを判定する(220)。

【0061】グループ間連絡マスタでなければ、グループ間連絡メッセージ宛先NSO識別子527が示す他グループのネーミングサービスオブジェクトを宛先として、上記受信メッセージM2を転送する(222)。もし、グループ間連絡マスタに指定されていた場合は、メッセージシーケンス番号632を更新(221)した後、受信メッセージM2を転送する(222)。一定時間内に応答がなければ(223)、上記受信メッセージM2の転送動作を繰り返し、宛先のネーミングサービスオブジェクトから受信応答があれば、このルーチンを終了する。

【0062】尚、一定時間内に応答がなかった場合、図

17のステップ164～168と同様に、受信メッセージM2内のデータブロック配列に従って、宛先NSOを変えてメッセージの送信動作を繰り返すようにしてもよい。この場合、図11に示すグループ間連絡メッセージM2によれば、リスタートマスタのNSO識別子は、各グループの先頭の負荷情報ブロック635-1に設定されているため、グループ間連絡メッセージ宛先NSO識別子527を先頭ブロックに含むグループの次のグループから順次にリスタートマスタのNSO識別子を抽出して、メッセージM2の送信を繰り返す。上記メッセージM2内のデータブロック配列において、自NSO識別子が第1グループ以外のグループに所属している場合は、最終グループから先頭グループに戻って、リスタートマスタNSO識別子の抽出を繰り返し、自グループに至った時点で、リトライ動作を終了する。クラスタ構成のクライアントサーバシステムで、何れかのリスタートマスタが障害となった場合、一時的に負荷分散が乱れるが、予備リスタートマスタが動作するすれば、正常な負荷分散に復旧できる。

【0063】上述した本発明のクライアントサーバシステムによれば、複数のサーバマシンがネーミングサービス機能を備え、それぞれの負荷情報テーブルを定期的に更新できるようになっているため、各クライアント端末は、上記複数のサーバマシンのうちの何れかにネーミングサービス要求を行うことによって、ネーミングサービス自体の負荷を分散した上で、サーバマシンの動的負荷分散を実現できる。また、実施例で示したように、ネーミングサービス機能をもつ複数のオブジェクトがサーバマシンの負荷情報を互いに交換する1つの方法として、負荷情報連絡メッセージを複数のネーミングサービスオブジェクトに順次に循環させる方法を採用した場合、サーバ負荷分散制御に必要な負荷情報を比較的少ない通信量で効率的に収集、配布できるという利点がある。また、上記負荷情報連絡メッセージは、負荷情報のみならず、サーバの障害情報を同時に転送することができるため、障害サーバを除外したネーミングサービスも容易に実現できる。

【0064】尚、本発明のように、複数のネーミングサービスオブジェクトが並列的にネーミングサービスを実行した場合、一時的に負荷の最も軽いサーバマシンに新たなサービス割り当てが集中する可能性があるが、実施例で示したように、各サーバ毎にサービス割り当て回数をカウントし、カウント値が所定の閾値に達した場合は、割り当て対象から除外し、上記カウント値は負荷情報更新の都度クリアするようにしておけば、特定サーバへの負荷集中を回避したネーミングサービスが可能となる。

【0065】実施例では、説明の都合上、クラスタ内で転送される負荷情報連絡メッセージM1と、クラスタ間で転送されるグループ間連絡メッセージM2を異なるフ

ォーマットで示したが、上記負荷情報連絡メッセージM1のフォーマットは、クラスタ構成を採用する必要のない比較的小規模のクライアントサーバシステムに有効である。実施例で示したようなクラスタ構成のクライアントサーバシステムにおいては、各クラスタ内での負荷情報連絡とクラスタ間の負荷情報連絡の両方に、上記グループ間連絡メッセージM2と同一のメッセージフォーマットを適用できる。

【0066】実施例では、複数のサーバマシン間の論理的なループを形成し、負荷情報連絡メッセージM1を上記ループに沿って連続的に循環させたが、本発明の変形として、例えば、負荷情報連絡メッセージを、負荷情報収集のための第1メッセージと負荷情報配布のための第2メッセージに分けて定義し、第1メッセージがループを一巡した時点で、リスタートマスタが、受信された第1メッセージを第2メッセージとして再送信することによって、各ネーミングサービスオブジェクトに他サーバの負荷情報を通知するようにしてもよい。この場合、第2メッセージは、ループに沿って順次に転送する必要はなく、同報メッセージとして送信することによって、通信リンクに接続された全てのサーバマシンが同時に受信できるようにしてもよい。上記変形例によれば、第1メッセージを所定周期で間欠的に生成した場合でも、収集した負荷情報を第2メッセージで直ちに全ネーミングサービスオブジェクトに通知できるため、負荷情報連絡のための通信量を抑制して、全ネーミングサービスオブジェクトに実態に合致した負荷情報に配布でき、適切な負荷分散を実現できる。

【0067】

【発明の効果】以上の説明から明らかなように、本発明のクライアントサーバシステムおよびネーミングサービス方法によれば、サーバマシンにネーミングサービス機能を装備することによって、ネーミングサービスを分散して実行できるようにしているため、クライアント端末の台数が増加した場合でも、迅速なネーミング応答が可能となる。

【図面の簡単な説明】

【図1】本発明のクライアントサーバシステムの構成を示す図。

【図2】負荷情報連絡メッセージM1のフォーマットの1例を示す図。

【図3】負荷情報連絡メッセージM1の転送制御を説明するための図。

【図4】クラスタ構成のサーバシステムにおける負荷情報連絡メッセージM1とグループ間連絡メッセージM2の転送制御を説明するための図。

【図5】本発明のサーバマシンが備えるプログラムの構成を示す図。

【図6】負荷情報テーブル51の構成を示す図。

【図7】イニシャライズ定義ファイル52の構成を示す

図。

【図8】再エントリ要求バッファ53の構成を示す図。

【図9】サーバオブジェクト位置情報テーブル54の構成を示す図。

【図10】ネーミングサービス要求メッセージのフォーマットを示す図。

【図11】グループ間連絡メッセージM2のフォーマットを示す図。

【図12】再エントリ要求メッセージM3のフォーマットを示す図。

【図13】動作可否確認メッセージM4とグループ間連絡宛先変更メッセージM5のフォーマットを示す図。

【図14】イニシャライズ処理ルーチン501を示すフローチャート。

【図15】振り分け処理ルーチン502を示すフローチャート。

【図16】時間監視処理ルーチン506を示すフローチャート。

【図17】負荷情報連絡処理ルーチン505を示すフローチャート。

【図18】ネーミングサービス処理ルーチン503を示すフローチャート。

【図19】時間起動処理ルーチン504を示すフローチャート。

ャート。

【図20】再エントリ処理ルーチン508を示すフローチャート。

【図21】グループ間連絡処理ルーチン507を示すフローチャート。

【符号の説明】

1：クライアント端末、2：サーバマシン、3：通信リンク、4：サーバプログラム、40：サーバオブジェクト、

5：ネーミングサービスプログラム、50：ネーミングサーバオブジェクト、

51：負荷情報テーブル、52：イニシャライズ定義テーブル、

53：再エントリ定義バッファ、54：サーバオブジェクト位置情報テーブル、

M1：負荷情報連絡メッセージ、M2：グループ間連絡メッセージ、

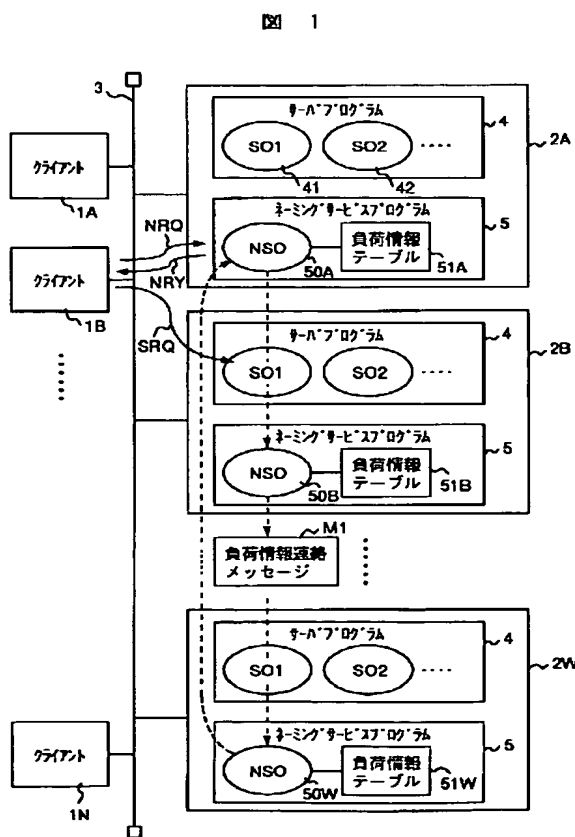
M3：再エントリ要求メッセージ、M4：動作可否確認メッセージ、

M5：グループ間連絡宛先変更メッセージ、

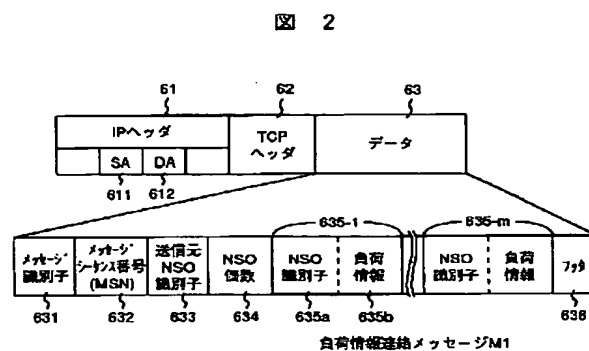
NRQ：ネーミングサービス要求メッセージ、NRY：ネーミングメッセージ、

SRQ：サービス処理要求。

【図1】



【図2】



【図7】

図7

イニシャライズ定義ファイル: 52

521	リストマシナ表示フラグ	グループ間連絡マシナ表示フラグ	525
522	リストマシナNSO識別子	グループ識別子	526
523	予備リストマシナ表示フラグ	グループ間連絡メッセージ宛先NSO識別子	527
524	次NSO識別子	グループ間連絡メッセージ送信元NSO識別子	528

【图 1 1】

11

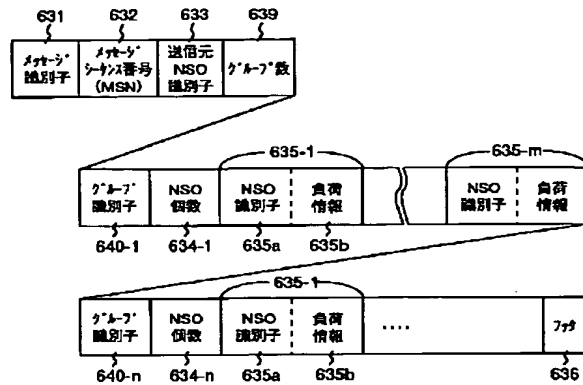
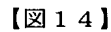
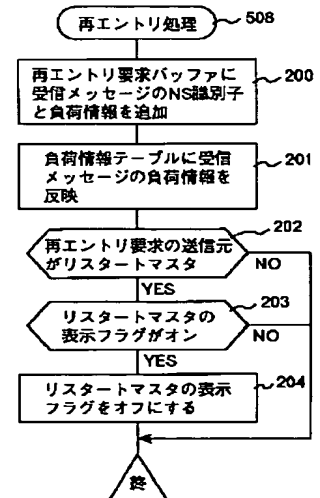


图 12

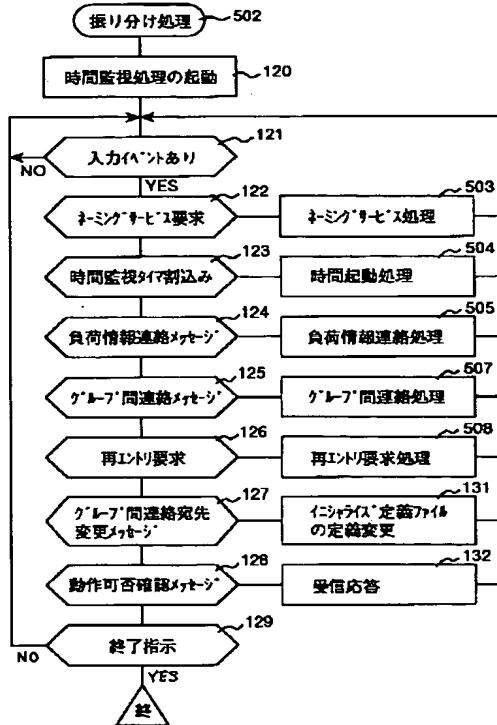


【図 20】



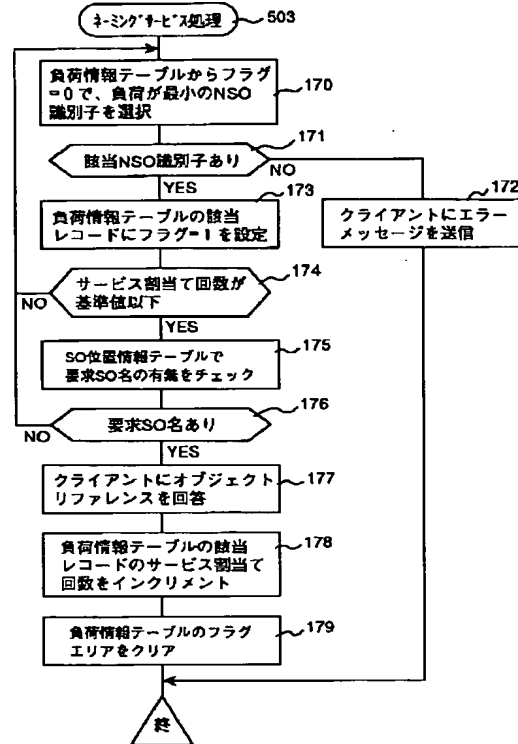
【図15】

図 15



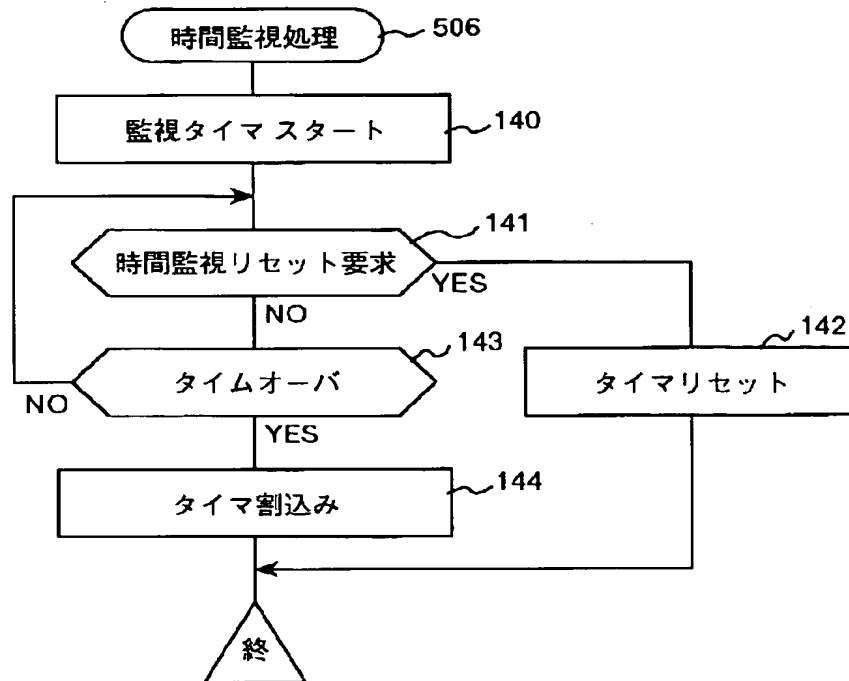
【図18】

図 18



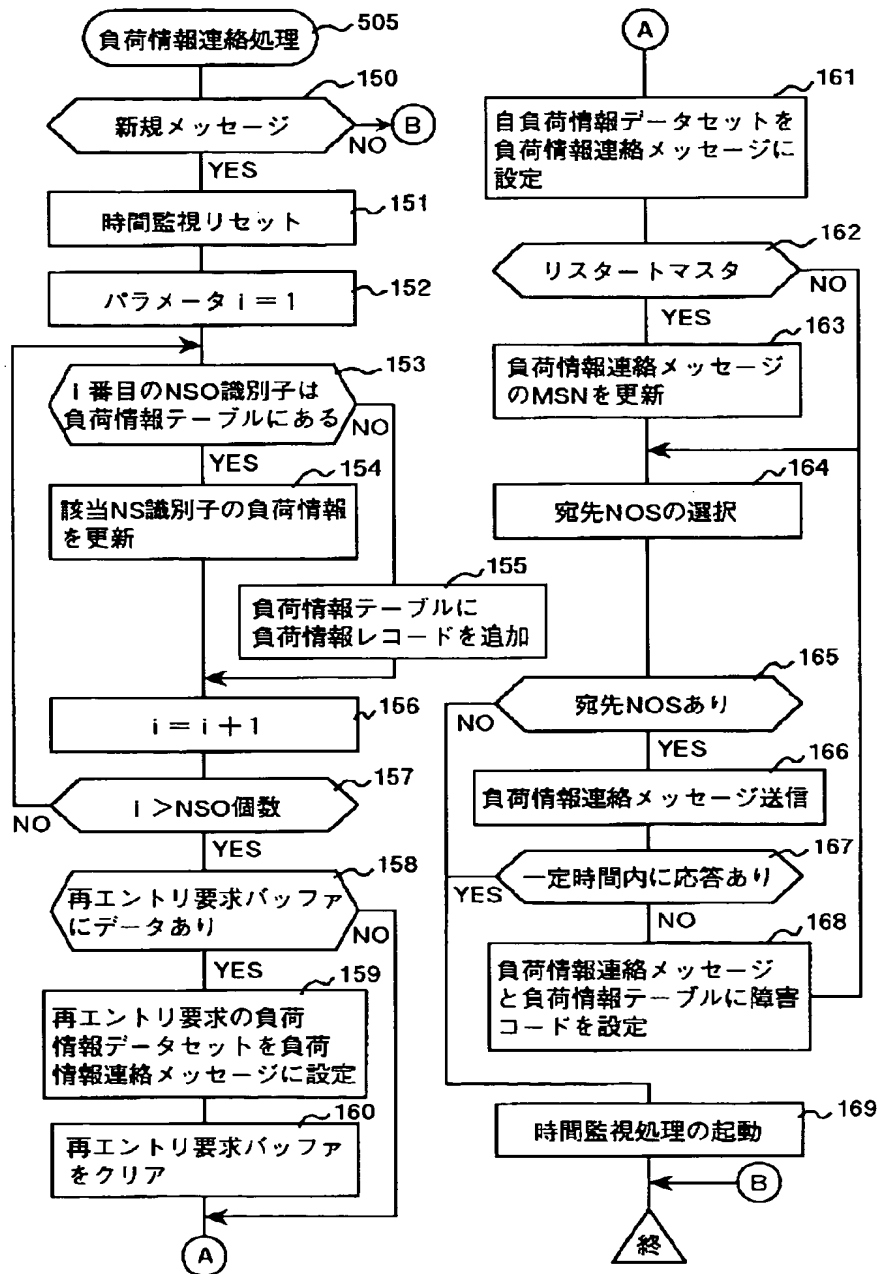
【図 16】

図 16



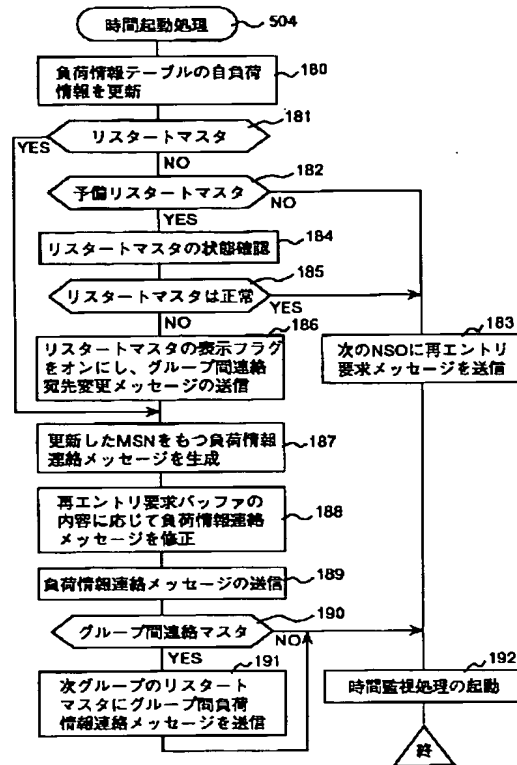
【図17】

図 17



【図19】

図 19



【図 21】

図 21

